



Contribution ID: 1

Type: not specified

”Improving” Prediction of Human Behavior Using Behavior Modification

Monday, 28 September 2020 15:15 (45 minutes)

The fields of machine learning and statistics have invested great efforts into designing algorithms, models, and approaches that better predict future observations. Larger and richer data have also been shown to improve predictive power. This is especially true in the world of human behavioral big data, as is evident from recent advances in behavioral prediction technology. Large internet platforms that collect behavioral big data predict user behavior for their internal commercial purposes as well as for third parties, such as advertisers, insurers, security forces, and political consulting firms, who utilize the predictions for user-level personalization, targeting and other decision-making. While machine learning algorithmic and data efforts are directed at improving predicted values, the internet platforms can minimize prediction error by “pushing” users’ actions towards their predicted values using behavior modification techniques. The better the internet platform is able to make users conform to their predicted outcomes, the more it can boast both its predictive accuracy and its ability to induce behavior change. Hence, internet platforms have a strong incentive to “make the prediction true”, that is, demonstrate small prediction error. This strategy is absent from the machine learning and statistics literature. Investigating the properties of this strategy requires incorporating causal terminology and notation into the correlation-based predictive environment. However, such an integration is currently lacking. To tackle this void, we integrate Pearl’s causal $do(\cdot)$ operator to represent and integrate intentional behavior modification into the correlation-based predictive framework. We then derive the expected prediction error given behavior modification, and identify the components impacting predictive power. Our formulation and derivation make transparent the impact and implications of such behavior modification to data scientists, internet platforms and their clients, and importantly, to the humans whose behavior is manipulated. Behavior modification can make users’ behavior not only more predictable but also more homogeneous; yet this apparent predictability is not guaranteed to generalize when the predictions are used by platform clients outside of the platform environment. Outcomes pushed towards their predictions can also be at odds with the client’s intention, and harmful to the manipulated users.

Presenter: SHMUELI, Galit (Editor of INFORMS Journal on Data Science and Tsing Hua Distinguished Professor at National Tsing Hua University, Taiwan)

Session Classification: Invited Plenary Talk